Probabilistic Constraint Satisfaction with Structural Models: Application to Organ Modeling by Radial Contours

Russ B. Altman, MD, PhD
Section on Medical Informatics
Stanford University
altman@camis.stanford.edu

James F. Brinkley, MD, PhD
Department of Biological Structure
University of Washington
brinkley@cs.washington.edu

One of the key challenges within medical information sciences is the development of useful models for biological structure and its variability. Many biomedical problems involve the elucidation of structure (for example, from experimental data or from imaging studies), and structural models can often drive the process of inferring precise structure from data. Ideally, model-driven data interpretation combines knowledge about the generic features of a class of biological structures (as contained within a model) with data that provide specific information (often noisy) about a particular instance of the class. In this paper we briefly discuss model-driven determination of biological structure as an example of a structural constraint satisfaction problem. We describe a probabilistic implementation of structural constraint satisfaction, and show that our formulation of a particular organ modeling technology (Radial Contour Models) exhibits promising performance. Our results demonstrate the utility of probabilistic models for the solution of structural constraint satisfaction problems.

INTRODUCTION

Reasoning about structure often involves moving from generic models (based on a population of training examples) to specific models (based on a combination of generic models and case-specific data). Such tasks can be mapped onto a class of problems known as *geometric constraint satisfaction problems* [9]: solutions are sought which are compatible with the model and satisfy the case-specific constraints. Constraint satisfaction problems can be formulated as *constraint graphs*: each parameter (or *variable*) to be estimated is a node within a constraint graph, and has an associated set of possible values (the *domain* for the node). The *arcs* within the graph are the constraints between (or among) different variables. Starting with a long initial list of possible values, the goal is to eliminate incompatible values, and arrive at a small number of values consistent with the constraints. When all value lists have been maximally pruned, then the set of possible individual solutions can be generated by serially visiting all the nodes, and selecting from each node a value that is compatible with the values selected for previously visited nodes. The theory of constraint satisfaction networks is described in detail in [9].

We have previously shown that geometric constraint satisfaction problems can be solved using a formulation in which the values associated with each node take on discrete values, and the constraints are used to prune these lists [4,6]. In addition, we have shown that certain constraint satisfaction problems can be formulated probabilistically so that each node is represented as a continuous distribution [1,2]. Constraint arcs are then represented as probability distributions between the nodes. The constraint distributions, then, modify the node distributions using the laws of conditional probability. This formulation has been shown to work well in the case of protein structure interpretation from NMR data [3].

In this paper, we investigate the utility of the probabilistic reformulation of constraint satisfaction problems in the context of a different biological structure problem: organ contour modeling for analysis of CT scans. We have previously reported a formalism for modeling organ cross-sectional shapes in CT scans, called the Radial Contour Model (RCM) —an interval based geometric constraint network [6]. The organ cross-section can be represented by a centrally located point from which a set of radial "spokes" emanate from 0 to 360 degrees (at some increment). The contour of an organ can be described as a set of distances along each of the spokes. A generic model describing the shape and range of variation of a shape class is built by examining a training set of similarly-shaped organ cross-sections and representing constraints as the range of slopes of the lines between neighboring radials. In terms of constraint satisfaction, therefore, the variables (nodes) are the distances along each of the radials. The set of values for each node initially can range from 0 to infinity, but based on training data can be narrowed to a smaller range. The arcs between nodes are the constraints placed on the radials by their expected relationship. We have shown that this representation is useful for semi-automated segmenting as well as organ identification [6].

Using the same constraint satisfaction analogy, we have defined a probabilistic model in which the variables are also radial distances, but are represented as continuous probability distributions over the distances (instead of discrete ranges). We have defined

constraints based on the covariation of these variables (instead of ranges of ratios). We show that as we introduce data about a particular radial distance into our model (as would occur, for example, in a model-based image segmentation system), we can update our estimate of the value of the other radials using the laws of conditional probability. Specifically, we use a Bayesian parameter estimator [8]. Our estimates improve rapidly and converge on the actual contour relatively quickly. We will describe the machinery of our model, and illustrate preliminary tests of its performance. Our work differs from previous work in probabilistic constraint satisfaction [7] because we use continuous variables, instead of assigning probabilities to discrete variable values.

METHODS

We represent a structural model with two main data structures: a vector of the parameters to be estimated (which represents the mean value of our variable or node), and a variance/covariance matrix which represents the range of values that the parameters can assume (the variance of each parameter) as well as the constraints between parameters (the covariance between parameters). In the case of the RCM, the state vector is a list of the average radial distance from the origin for each of the angles from 0 to 360. In the work described here (as in Brinkley [6]), we always use 24 radials separated by 15 degrees each. Thus, our state-vector has 24 elements:

$$x = \begin{bmatrix} r0 & r1 & r2 & ... & r22 & r23 \end{bmatrix} \quad (1)$$

The second element of our representation is the variance/covariance matrix which captures two essential features of a structural model, (1) the variance over the training set of each of the elements of the state vector (that is, the variance of each ri) is contained along the diagonals, and (2) the covariance between radials, a linear approximation of the dependence between two radials, is contained along the off-diagonals. The variance/covariance matrix, therefore contains 24 x 24 = 576 elements:

$$C(x) = \begin{bmatrix} \sigma^2_{r0} & \sigma_{r0r1} & \cdots & \sigma_{r0r23} \\ \sigma_{r1r0} & \sigma^2_{r1} & & \cdot \\ \cdot & \cdot & \cdot & \\ \sigma_{r23r0} & \sigma_{r0r1} & \cdots & \sigma^2_{r23} \end{bmatrix} \quad (2)$$

where variance is defined as:

$$\sigma^2_{ri} = E(ri^2) - E(ri)^2 \quad (3)$$

and covariance is defined as:

$$\sigma^2_{rirj} = E(rirj) - E(ri)E(rj) \quad (4)$$

The covariance between two radials is a linear estimate of how the radials vary with one another. A positive value for covariance indicates that two radials are positively correlated, whereas a negative value for covariance indicates that the two radials are negatively correlated. (Note that covariance is the numerator of the standard correlation coefficient, which is normalized by the product of the variances).

Given a training set, we can calculate a model by calculating the empirical means (to provide the average values for the state vector, x), and calculating the empirical variances and covariances (using the above formula) in order to instantiate the matrix C(x). The resulting data structures are a probabilistic representation of a constraint network. The constraints between different parameters are represented as a covariance, and there is a covariance between every pair of parameters. Thus, the covariance matrix is a dense network of interconnections that describe the dependence of each parameter on the others. The covariances are linear estimates of the relationship between two parameters. They create a highly connected data structure in which multiple dependencies can be propagated over many variables.

Introducing Data and Refining the Model

The model calculated in a manner described above is a generic model that summarizes the range of structures in the training set (see Figure 2A). As information is gathered about a particular structure, the model can be used to create a customized estimate of the structure, incorporating the measured data as well as the a priori model (Figures 2B, 2C, and 2D). Measurements are modeled as a function of the parameter vector, x, and a mean-zero Gaussian noise term, v:

$$z = h(x) + v \quad (5)$$

where z is the measured value, h(x) is the function describing how z would be predicted from the state vector, and v is a random noise process that causes z and h(x) to be different. In general, z and v are vectors, and h(x) is a vector function. In the case of the RCM model, the measurements are the radial distance of an organ edge from the origin of the model. Thus, in this case,

$$h(x) = ri \quad (6)$$

where ri is the radial being measured. V is a function of the measurement technology; in this work we assume that radials are determined to within 1.0

mm. v is assumed to be normally distributed (often a good assumption in the case of biological structure). We have shown elsewhere that $h(x)$ can be non-linear and still yield good solutions using iterative techniques [1].

Given a model (x and $C(x)$) and a measurement, z, our goal is to create an updated model that incorporates the measurement and its error into the model. The equations for combining these data are derived from Bayes' rule [8]:

$$x(new) = x(old) + K [z - h(x(old))] \qquad (7)$$

$$C(new) = C(old) - K H C(old) \qquad (8)$$
where
$$K = C(old) H^T [H C(old) H^T + v]^{-1} \qquad (9)$$

and

$$H = \frac{\delta h(x)}{\delta x}\bigg|_x \qquad (10)$$

These updating formulae are Bayesian in the sense that they combine a prior information contained in the model ($x(old)$, $C(x(old))$) with measurement data and measurement noise (z, v) to provide updated estimates. They can be replaced with equivalent iterative equations in the case where $h(x)$ is non-linear [8].

Experiments with RCM

In order to test this representation on biological organ data, we used the data described in [6] for one view of the kidney and one view of the spleen. These were chosen because they are irregular contours. The kidney at the level of the hila has a marked notch, and the spleen has an amoeboid shape at the cut level used.

For each of the two models, we used a training set of 15 cross sectional CT scans, which were manually aligned and for which radii were measured manually. We tested the algorithm on three test cases, which were not used for training. For each test case, we introduced the radii at 0 and 180 degrees (the "poles"), and then introduced the additional radial constraints sequentially. This procedure simulates that used previously, in which the model was sequentially updated by edge data obtained along selected radials in the image [6]. Each radii was assumed to be introduced with a variance of 0.06 mm^2, which implies two standard deviations of 0.5 mm on each side of the mean. In order to test the suitability of the model and computational machinery, we measured three things:

1) The initial error of the test cases. We would expect that a good model would describe cases *not included in the training set* as well as it describes those included in the training set--this is a measure of the *a priori* validity of the model.

2) The distance of each radial estimate from the solution. We would expect that as data are introduced the average error would decrease. Clearly, if all the distances are given (with low variance), the solution is specified. It is also important to see how accurate the estimate is when only a few distances are provided.

3) The average uncertainty of each parameter as a function of the number of distances introduced. We would expect that the model's uncertainty would decrease monotonically as information is introduced.

RESULTS AND DISCUSSION

The generic model of the kidney hila resulting from the 15 test cases is illustrated in Figure 2A. The remainder of Figures 2 shows progress of the algorithm as it refines its estimate of the organ for which measurements are being provided. There are three ways to judge the initial adequacy of our model.

1. *Initial error of test data.* We evaluated how well the model describes test data which it has not yet seen. For each radii in a test case, we calculated the number of standard deviations it was from the mean of the *a priori* model. This is calculated as

$$E_c = \frac{(r_{actual} - r_{model})}{\sqrt{v_c}} \qquad (11)$$

where E_c is the error to be calculated for a given radial, r_{model} is the radius in the calculated model, v is the variance of that radial within the model, and r_{actual} is the actual radius of the test case. This formula provides the error in units of standard deviations. When we evaluated our test cases against the generic model, all the radii occur within 2.0 standard deviations of the expected values and expected variance. Thus, the model captures the generic shape of the organs adequately.

2. *Rate of convergence of generic model to specific model.* We introduced the radial distances (and their uncertainty) into the model using equations (7) and (8) above. After each radius was introduced, we had a more customized model that combines the initial generic model with the measurement data. Figure 2 shows graphically how the model is refined as more constraints are introduced. Figure 1 shows how the

494

RMS difference between predicted radii and actual radii is reduced as constraints are introduced. It is not surprising that the model is a good one after all 24 radial distances are supplied, but it is somewhat surprising that after only 4 radial distances are provided (in the case of the kidney), the algorithm has come within 1.8 mm of the 24 correct radii (from a starting RMS error of 5.5 mm). This illustrates the power of using the covariance information.

Figure 1. Average RMS error for test cases as function of number of constraints introduced. The spleen data is noisier, and thus starts with greater uncertainty.
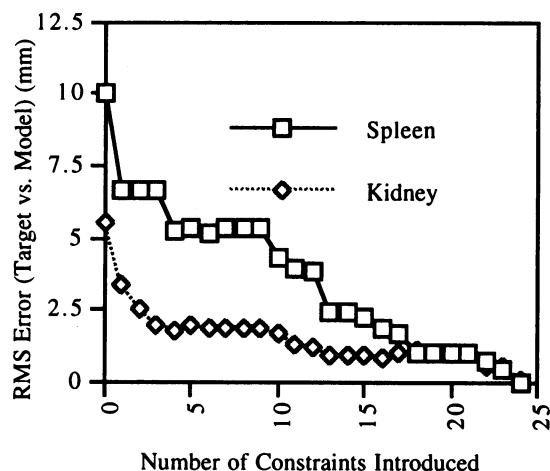

Number of Constraints Introduced

Table I. Average radial variances for 3 test cases as constraints are introduced sequentially.

Model	Kidney (mm^2)	Spleen (mm^2)
Initial Average Var.	44.50	90.25
Avg. Var. after 2 radii	28.60	52.34
Avg. Var. after 8 radii	2.84	16.13
Avg. Var. after 16 radii	0.20	3.10
Avg. Var. after 24 radii	0.06	0.06

3. *Rate of variance reduction.* Table I shows how the average uncertainty in radial distances decreases as information is supplied. In the case of the kidney, there is an initial uncertainty for each radius of about 45 mm^2. The introduction of 8 constraints reduces this uncertainty to 2.8 mm^2. The introduction of all 24 constraints reduces the uncertainty to within the tolerance of the measurements. Table I also shows similar information for the spleen data, which is

considerably more noisy. As can be seen, the initial variance is much larger, and more radial constraints are required to reach the same level of certainty as with the kidney

CONCLUSIONS

The results reported here show the applicability of the probabilistic formulation of constraint satisfaction to modeling organs. We have demonstrated its performance in the case of RCMs. Our model is general and can be applied directly to other parametric organ models using equations 1-10. We have shown that our model is easy to build from a training set using the standard definitions of mean, variance and covariance. We have shown that our model is able to express the average structure of two irregular organ contours (spleen and kidney) as well as their variability. Finally, we have shown that the update equations (7-10) are useful for introducing constraints and producing updated estimates of structure (by combining the constraints with the pre-existing model in a Bayesian fashion). We have not demonstrated that our probabilistic formulation is superior to the discrete approaches previously described. However, the performance on these test data sets (and its performance in the realm of molecular structure elucidation [3,4]) indicates that it may provide a powerful alternative.

REFERENCES

[1] Altman, R. and O. Jardetzky, *The Heuristic Refinement Method for the Determination of the Solution Structure of Proteins from NMR Data*, in *Nuclear Magnetic Resonance, Part B: Structure and Mechanisms*, N.J. Oppenheimer and T.L. James, Editor. 1989, Academic Press: New York. p. 177-218.

[2] Altman, R.B., R. Pachter, and O. Jardetzky, *The Determination of Structural Uncertainty from NMR and Other Data: the lac repressor headpiece*, in *Protein Structure and Engineeing*, O. Jardetzky, Editor. 1989, Plenum Press: New York. p. 79-95.

[3] Arrowsmith, C., R. Pachter, R. Altman, and O. Jardetzky, *The Solution Structures of E. coli trp Repressor and trp Aporepressor at an Intermediate Resolution.* European Journal of Biochemistry, 1991. **202**: p. 53-66.

[4] Brinkley, J.F., R.B. Altman, B.S. Duncan, B.G. Buchanan, and O. Jardetzky, *The Heuristic Refinement Method for the Derivation of Protein Solution Structures: Validation on Cytochrome-b562.* Journal of Chemical Information and Computer Science, 1988. **28**(4): p. 194-210.

[5] Brinkley, J.F. *Hierarchical geometric constraint networks as a representation for spatial structural knowledge.* 16th Annual Symposium on Computer Applications in Medical Care, Nov. 8-11, 1992, pp 140-144 (1992).

[6] Brinkley, J.F. *A flexible, generic model for anatomic shape: application to interactive two-dimensional medical image segmentation and matching.* Comp. Biomed. Res. In Press. 1993.

[7] Faber, T.L., Stokely, E.M., Peshock, R.M. and Corbett, J.R. *A model-based four-dimensional left ventricular surface detector.* IEEE Trans. Med. Imaging. 10, 321 (1991).

[8] Gelb, A., *Applied Optimal Estimation.* 1984, Cambridge, Massachusetts: MIT Press.

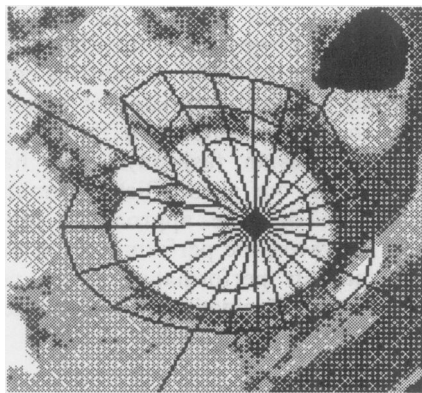[9] Mackworth, A.K. *Consistency in networks of relations.* Artificial Intelligence. 8, 99 (1977).

Figure 2A. The generic probabilistic model shown superimposed on the actual CT cross-section of a test-case kidney. The mean model is shown as the center contour. Contours at 2 SD above and below the means are also drawn. This represents the initial model as generated from the 15 test cases.
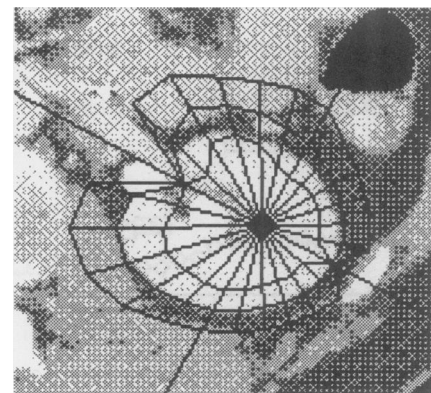

Figure 2B. The model after two poles (at 0 and 180 degrees) have been introduced. The mean contour and the errors are updated to reflect the new information. There are still significant deviations of the model from the actual contour--two radii (and the model) do not fully specify the contour.
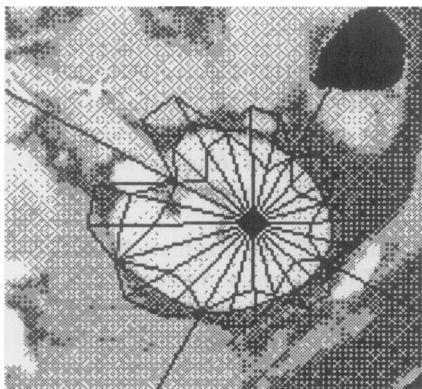

Figure 2C. The model after introduction of 8 constraints. The mean model now matches the actual contour well. The errors are small along radials for which constraints have been introduced. The errors are larger in adjacent areas for which there are no measurements, although the mean estimate is good.
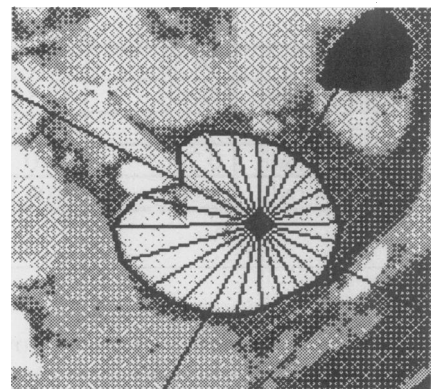

Figure 2D. After all 24 radial distances are provided, the model converges on the exact contour, with very low uncertainty (mean and 2 SD errors virtually coincide).